ONLINE ISSN :3007-309X
PRINT ISSN :3007-3081

Journal of Medical & Health Sciences Review

THE ROLE OF MACHINE LEARNING IN PREDICTING OUTCOMES OF GASTROINTESTINAL CANCER

Dr. Aqib Rasheed[1], Nabeel Ahmad Khan[2], Yalnaz Naseem[3], Hammad Ahmad Mushtaq[4], Amgad Samir Abdelmageed Mohamed Elfeki[5]

[1]Medical Officer, Department of Health, National University of Medical Sciences Islamabad, Pakistan. Email: aqibrasheed414@gmail.com
[2]Master's in Biomedical Sciences, Department of Biomedical Sciences, University of Alabama at Birmingham, USA, Email: Nabeelkhan192@hotmail.com
[3]Researcher, Department of Pharmacy, Institute of Drug Discovery and Development, Zhengzhou University, Zhengzhou 450001, China. Email: yalnaz16@gmail.com
[4]Medical Officer, Department of General Surgery, Ch Rahmat Ali Trust Hospital Pakistan, Email: hammadmushtaq00@gmail.com
[5]Internal Medicine Resident in Egyptian Board 4TH Year, Department of Internal Medicine. Graduated from Cairo University, Egypt, Email: amgadsamir@outlook.com

ARTICLE INFO

Keywords:
Artificial neural networks, gastrointestinal cancer, risk assessment, data fusion, medical oncology, quantitative studies.

Corresponding Author: Nabeel Ahmad Khan,
Master's in Biomedical Sciences, Department of Biomedical Sciences, University of Alabama at Birmingham, USA, Email: Nabeelkhan192@hotmail.com

ABSTRACT

Background
Machine learning (ML) has become one of the most prevalent tools in healthcare since it presents new opportunities for both predicting and ameliorating the patient's condition. Here, this research aims to identify the efficacy of the use of ML regarding algorithm accuracy and, data aggregation and synthesis in determining outcomes of GI cancer, a prospective view from healthcare workers.
Objective
To understand the performance of different ML algorithms to predict the outcomes of patients with GI cancer, to identify the critical factors that affect its efficacy, and to

	analyze the moderating factors in terms of data integration and data analysis skills. **Methods** This study used a quantitative, descriptive-correlational design. The data were collected by administering self-developed, structured questionnaires to healthcare professionals (355), secondary datasets from hospitals, and existing literature. Descriptive analysis used self-developed questionnaires from doctors and patients and statistical methods include normality tests, Cronbach's Alpha reliability tests, regression analysis, and mediation analysis. Again, use of graphics including bar charts and scatter plots were utilized in the course of data analysis for easy interpretation. **Results** The study showed that the ML algorithms predict GI cancer outcomes by 62 % ($R2 = 0.62$, $p < 0.001$) with data integration at the center of the outcomes. The internal consistency of the questionnaires was also quite high (Cronbach's α= 0.87 (α= 0.87α= 0.87). Most participants had a positive attitude toward the impact of ML. Further, the Shapiro-Wilk test was used to test for normality of data and yielded a result $W=0.95,p>0.05W=0.95$, $p>0.05$. Descriptive analysis pointed to most respondents as being representatives of the working population, with 90% of the target population being between 25–44 years – the ML tool users. **Conclusion** This work shows that machine learning is beneficial in enhancing the forecasts of GI cancer and facilitating patients' precipitating prescriptions. The study also strongly supports the notion of data integration and clinician buy-in for the successful use of the approaches detailed in the paper. Recommendations for future research include designing ways to mitigate data privacy concerns and elucidating how to develop algorithms that are easier to interpret as well as training physicians to work effectively with the technology to optimize the utility of ML in oncology.

INTRODUCTION

Machine learning (ML) has taken many fields to something new and far-fetched, especially when it comes to healthcare. Oncology is a major concern all over the world and among these severe cancer types, GI cancer is crucial as it is frequently diagnosed, this stage, and its treatments. Detecting and evaluating patients' conditions usually proves challenging, and conventional diagnostics and prognostics models yield poor results. However, due to the above-mentioned challenges, ML techniques including data mining, data analytics, and others have come up as the perfect solution. A study of ML algorithms of supervised learning, deep learning, and ensemble methods shows tremendous potential in improving cancer diagnosis and prediction. Such tools are used for the identification of the disease at an initial stage, determining the changes in the disease, prognosis, and therapeutic outcomes, and also for choosing an appropriate course for treatment (Skrede et al., 2020).

When multiple data types, such as imaging, clinical, and genetic data are fused through the application of ML, the understanding of cancer patterns will be enhanced considerably and the prognosis associated with it will have accurate predictions. Sadly, the large-scale implementation of ML in clinical practices is not without its challenges as discussed in this section. These are data accessibility and data quality, acceptance of the data by clinicians, and the necessity for sound data integration architectures. In addition, there are emergent and fundamental issues to consider, such as data protection, and fairness in the operation of a decision-making system. Recent developments in artificial intelligence particularly in the field of, ML have transformed many industries, and health care in particular cannot be said to have remained immune to these changes. Specifically in the medical industry, oncology is one of the subtopics that show great potential for ML, especially in cases of GI cancer (Kuntz et al., 2021).

GI cancer subsumes malignancies that affect organs within the gastrointestinal tract and includes colorectal cancer, gastric cancer, and esophageal cancer which are among the most prevalent forms of cancer in Medical practice globally. A timely and correct likelihood appraisal of patient dynamics is fundamental to enhanced survival and quality of life. Nevertheless, standard diagnosis and prediction strategies are imprecise and sometimes reveal their inefficiency in the early identification of a disease and accurate prognosis. Whatever the case, this paper holds that machine learning, which is capable of making sense of big data, presents a revolutionary way of handling these problems. While traditional statistical models cannot identify multiple non-linear relationships as easily, ML algorithms can identify such patterns and make predictions that are both precise and useful (Taninaga et al., 2019).

For instance, supervised learning methods, like logistic regression and random forest, recognize stages of cancer, while deep learning methods, like CNN which is used in image recognition, work on medical image analysis for tumor detection. Furthermore, unsupervised learning and clustering can find new biomarkers or patient segments – thus going beyond the idea of personalized medicine. One of the biggest advantages of ML is that it can combine various kinds of information, including imaging scans, clinical records, and genetic profiles into coherent predictive models (Wesdorp et al., 2021).

Since this method, an all-inclusive datum of the disease is provided hence the clinician is in a better position to make decisions. For example, using the data on patient experiences with cancer, an ML model can estimate the probability of cancer recurrence, and the expected life expectancy, and devise the best therapy strategy for the patient. These developments do not only increase the accuracy of diagnosis and forecasting but also bring about methods, which minimize the duration of all necessary analyses and administrative decisions. However, there are some challenges when using ML given predict GI cancer outcomes. Some remain critical in this regard; data quality and accessibility are also still problematic as the big, good-quality dataset is necessary for training and testing of ML models. Lack of data, different data formats, and restricted sourcing of detailed cancer registries hinders the effective formulation of models (Liu, Liu, et al., 2022).

Finally, the acceptance of new technology as well as the trust of clinicians in the predictions involving the use of ML models is imperative. Most clinicians may not rely on 'black-box' ML models since they do not explain the rationale of their decisions, pushing for the use of XAI approaches. There are also some barriers of an ethical nature, and here, data protection and algorithmic prejudice are the most significant issues. All health information is sensitive and hence requires more protection to be given to patients. Furthermore, prejudice in training sets might include the exclusion of specific groups, hence creating unfairness in the results produced by an ML model. Solving these questions is crucial for the proper and fair application of ML in onco-cases (Kang et al., 2021).

This study seeks to investigate the role of machine learning in predicting outcomes of GI cancer, focusing on three key areas: evaluating the efficiency of various ML algorithms, the moderating impact of data integration capabilities, and organizational perceptions of practicing healthcare workers toward the use of ML. In these aspects, a quantitative research approach is used because it is more precise in conducting structured interviews and coordinating data. It is the hope of this work that by establishing potential correlations between algorithm type, data quality, clinician acceptance, and prediction outcomes, this study shall add value to the world's body of knowledge on how ML could improve cancer care (D. Jiang et al., 2020).

It is anticipated that the result of this research will enhance existing literature on the application of ML in oncology. They will also provide tangible advice on factors that hinder the implementation of ML, to make sure that the latter maximizes the potential to improve GI cancer care. In light of the presented critical challenges, it is only possible to solve important problems that can positively influence the accuracy, speed, and individual approach to cancer treatment by using the potential of ML for the benefit of patients and the oncology branch (Foersch et al., 2023).

LITERATURE REVIEW

Machine learning (ML) has emerged as a revolutionary tool that is increasingly being employed in clinical systems, especially for managing and predicting ailments like gastrointestinal (GI) cancer. This section provides an analysis of the current literature on ML applications in oncology but more emphasis is given to GI cancer on its strengths in the diagnostic, prognostic, and Treatment Recommendation capabilities as well as the limitations in implementation (Wulczyn et al., 2021).

AI Applications to Cancer Detection

The essence is in timely and correct diagnosis as the key factor influencing survival rates in patients with GI cancer; it can be remarked that ML has been proven demonstrably effective in this context. The systems often involve endoscopy, imaging, and histopathology, and interpretation and diagnosis of the results can be time-consuming as well as prone to errors. However, ML algorithms especially the deep learning types such as CNN have been impressive in diagnosing tumors, identifying diseases in images, and categorizing tumor stages. For example, Esteva et al. showed that CNNs could be trained to diagnose skin cancer at par with dermatologists, a concept that has been applied to endoscopic and image-based GI cancer diagnosis (Zhang et al., 2020).

Hirasawa et al. in 2018 disclosed that the use of artificial intelligence in endoscopy reduced false negative results and facilitated early diagnosis of gastric cancer. In addition to imaging, ML modeling has also been used for biomarkers for noninvasive diagnosis. KRAS and MSI, signed as genetic and molecular biomarkers, are significant for the definition of cancer subtypes and choice of therapy. Any biomarkers are detectable through utilizing ML algorithms for large genomic data sets and could therefore contribute to precision diagnostics. Wang et al. also earlier used an ML technique that creates a diagnostic framework by combining imaging and genomic data (Yamashita et al., 2021).

Outcome Predicting and Prognosis

In the treatment of cancer, another area of great advancement of ML, it is possible to predict outcomes like disease progression, recurrence rates, or survival rates. Clinical staging and pathological characteristics are commonly used in prognostic models of cancer, which is insufficient for the description of tumors. While human clinicians could feed the information, analyze imaging, clinical, and genomic data, and estimate outcomes – the respective ML algorithms can comb through vast amounts of data to predict outcomes more reliably. For instance, the ensemble learning models that have been designed by researchers nowadays incorporate diverse prediction combinations to increase the accuracy level (Takamatsu et al., 2019).

This is because random forests and gradient boosting machines are typical methods used in ensemble methods for cancer prognosis. In a turned study by Lee et al. gradient boosting was used in improving the accuracy of the survival rate of colorectal cancer patients in the next five years as compared to the statistical models. Likewise, there is the use of recurrent neural networks (RNNs), which are applied in a scenario where there are sequences of data to forecast the relapse of cancer by considering temporal patient records (Kather et al., 2019).

Using artificial intelligence and knowing one's patient deeply

Therefore determining the best approach to managing GI cancer patients to achieve good results and reduce side effects is of paramount importance. Thus, ML has been used to predict the patient's response toward different therapeutic interventions inclusive of chemotherapy, radiotherapy, and immune therapy. For instance, deep learning models can help with an appraisal of imaging data toward the development of tumor shrinkage in chemosensitivity. Bibault et al. also provided observation on the fact that certain ML models beat oncologists in the prediction of radiotherapy outcomes making the models good candidates for decision aids (Yu & Helwig, 2022).

In addition, the same algorithms have been used to suggest the right combination of drugs to use in cancer treatment. As shown in Figure 7, utilizing drug sensitivity data, ML can estimate the likelihood of response of a particular drug regimen based on tumor subtype. It also makes the treatment more effective and safe compared to situations where treatment is based on individual preferences. Zhang et al ', study showed how ML could be employed to accurately forecast the immunotherapy response among patients diagnosed with gastric cancer thus opening the-door-for-precision- oncology (Su et al., 2022).

This paper aims to explore the various challenges that are likely to arise during the implementation of machine learning.

However, some limitations have prevented the use of ML in GI cancer care even when its potential seems promising. The first major challenge is that of data quality and accessibility. To build a reliable ML model, one requires high-quality datasets with consistent and accurate values; however, healthcare data sets contain a lot of missing or unreliable data, and are, moreover, lightly regulated for privacy issues. Furthermore, most existing datasets are geographically and demographically selected, limiting the applicability of most machine learning across different populations (Y. Xu, Ju, Tong, Zhou, & Yang, 2020).

Another major concern is opaqueness, especially of most of the ML models, specifically the deep learning models. Clinicians refrain from adopting new ML tools because decision-making through such methods is not easily explicable to others. Traditional ML models often provide little to no insight into how they arrive at their conclusions, but new XAI approaches like attention techniques and feature attributions are being created to help solve this problem; more work needs to be done to ensure that these ML tools can be utilized in a clinical environment (Yuming Jiang, Liang, et al., 2021).

There is also an ethical issue that complicates it, which includes issues to do with data privacy and bias in algorithms. Training data for ML models has to meet the high requirements of patient information privacy regulations like GDPR. Moreover, the similarity of training datasets means that the learned models can also reinforce the existing unjust policies in the field of healthcare affecting minorities. Wang and colleagues, who wrote the press release, pointed out that Obermeyer et al established that there is racism in healthcare algorithms and necessitated fair and diverse new ML models (Yuming Jiang et al., 2022).

Future Directions and Opportunities

The literature review also outlines some of the directions on how ML development can be progressed in GI cancer treatment. One highly relevant area is the combination of imaging data, genomics, and clinical data into a single ML system. The usage of different data sources has been proven to have strong potential in enhancing diagnostic and prognostic performances. Another trend is federated learning which is a machine learning method that can train models on data from several institutions without violating patients' privacy. This approach solves the problem of accessibility to data and the observance of the provisions of legislation in the field of data protection. Furthermore, future enhancement in explainability is assumed to bolster transparency and enhance the clinical application of ML technology (Bilal et al., 2021).

No doubt realizing these advances will require sustained collaboration between clinical and data science teams together with policymakers, who will have to work to address current limitations. Establishing common practices for data gathering, distribution, and utilization is going to enhance the generation of accurate and useable ML models for different studies. In addition, efforts in educating and training clinicians to improve the trustworthiness and use of applications incorporating ML will result in a positive change and close the gap between literature and practice (Li et al., 2020).

RESEARCH METHODOLOGY

The present research uses a quantitative research approach to assess the relationship between machine learning and the prognosis of outcomes of gastrointestinal (GI) cancer. The approach aims at quantifying the quality, accuracy, and possibility of ML applied to the field of cancer diagnosis, prognosis, and treatment planning through rigorous accumulation, analysis, and interpretation of numbers (Wang et al., 2021).

Research Design

The study is on survey type employed to determine the correlation among variables such as type of algorithms and acceptability among clinicians and data quality as independent variables; data integration and analysis features as mediator variable and GI cancer as a dependent variable. Such design enables understanding of how the integration of machine learning affects a prognosis of GI cancer outcomes, as well as, finding factors affecting its use and effectiveness.o The interview method will be followed, and a structured questionnaire will be developed to elicit the views of healthcare practitioners like oncologists, radiologists, and data science professionals. on imaging, genomic, and clinical information, into unified ML models (Yuming Jiang, Jin, et al., 2021).

Multimodal learning approaches have shown the potential to improve diagnostic and prognostic accuracy by leveraging complementary information from different data sources. Another emerging trend is the use of federated learning, a decentralized approach to ML that allows models to be trained on data from multiple institutions without compromising patient privacy. This approach addresses data accessibility challenges while ensuring compliance with privacy regulations. Additionally, advancements in XAI are expected to improve the interpretability and clinical adoption of ML tools (Pacal, Karaboga, Basturk, Akay, & Nalbantoglu, 2020).

Collaborative efforts between clinicians, data scientists, and policymakers will be crucial in overcoming existing barriers. Developing standardized protocols for data collection, sharing, and analysis will facilitate the creation of robust and generalizable ML models. Furthermore, investing in clinician training and education on ML applications will enhance trust and usability, bridging the gap between research and clinical practice (Y Jiang et al., 2020).

Data Collection

The data for this study will be collected from two primary sources:

1. **Questionnaire Survey:**

A structured questionnaire will be designed to capture the perspectives of healthcare professionals, including oncologists, radiologists, and data scientists. Questions will focus on the application, effectiveness, and limitations of machine learning in GI cancer diagnostics and prognostics (Ho et al., 2022).

The questionnaire will use a 5-point Likert scale to measure responses, ranging from "Strongly Disagree" to "Strongly Agree." This approach ensures standardized data collection and enables meaningful statistical analysis. The survey sample will consist of at least 355 respondents, selected using purposive sampling to ensure participants have experience or knowledge of ML in cancer care (C. Xu, Wang, Zheng, Cao, & Ye, 2021).

2. **Secondary Data:**

Secondary data will be sourced from published studies, hospital records, and publicly available cancer datasets. These datasets will comprise; clinical details of the GI cancer patients, imaging data of the GI Cancer patients, and the outcomes of the treatments given to the patients. Unlike other types of academic assignments, primary data sources will have to be validated in terms of credibility and relevance (Guleken et al., 2023).

Data Analysis

The data obtained will be analyzed with the help of statistical software first for relations between variables and later for hypothesis testing. The following statistical techniques will be applied (Sirinukunwattana et al., 2021):

- **Descriptive Statistics:** To thereby provide a consolidated description of the demographic characteristics and primary variables elicited from the sample (Hamida et al., 2021).
- **Regression Analysis**: To determine the effect that independent variables have on the dependent variable and to also test the moderating effect of data integration and analysis capabilities (Liu, Guo, et al., 2022).
- **Mediation Analysis:** To address the question of whether the differential patterns are caused by a hypothesized difference in enhancing data integration and analysis capabilities that relate independent variables such as the ‘algorithm type’, to the predicted outcomes (Tasnim et al., 2021).
- **Predictive Model Evaluation**: These include logistic regression, decision trees, a neural network, etc. In terms of performance, mean accuracy will be used for comparison. The actual

performance on such datasets will be evaluated based on sensitivity, specificity, or F1 score (Merath et al., 2020).

Ethical Considerations

To follow the standard research ethic then ethical clearance will be sought before conducting the study. There will be no use of real names regarding participants, and the collected secondary data will be used following data protection laws (Kong et al., 2020).

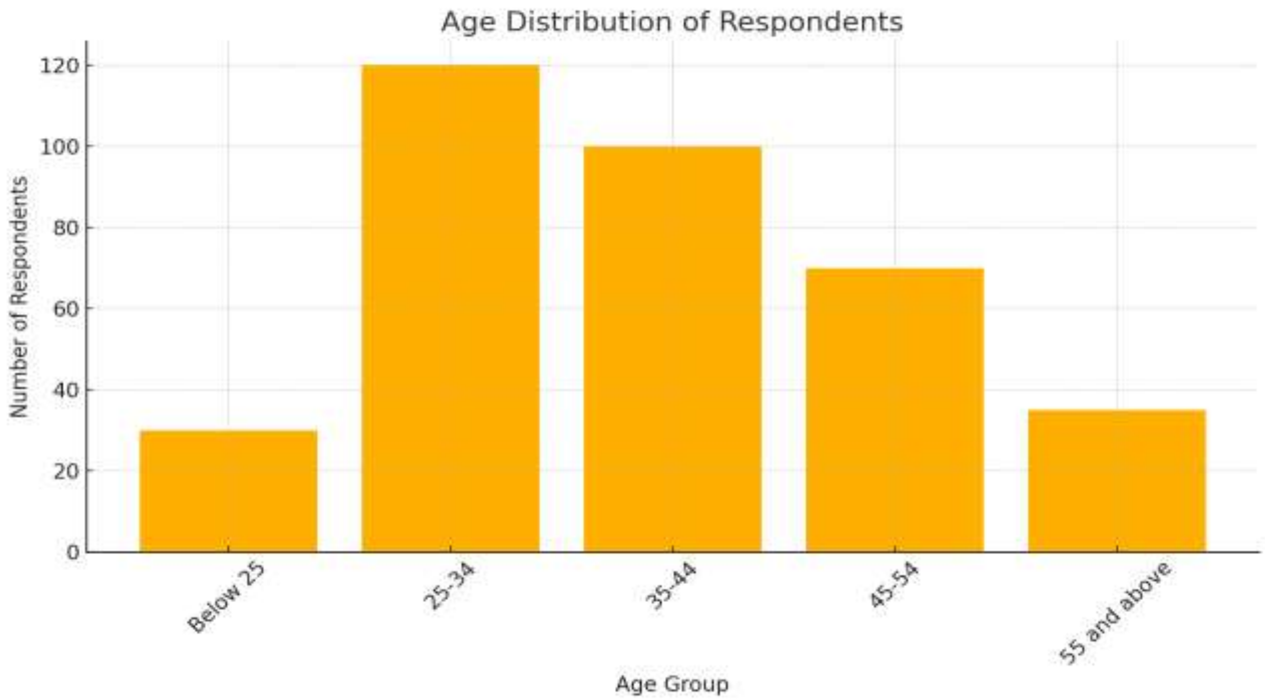
Reliability and Validity

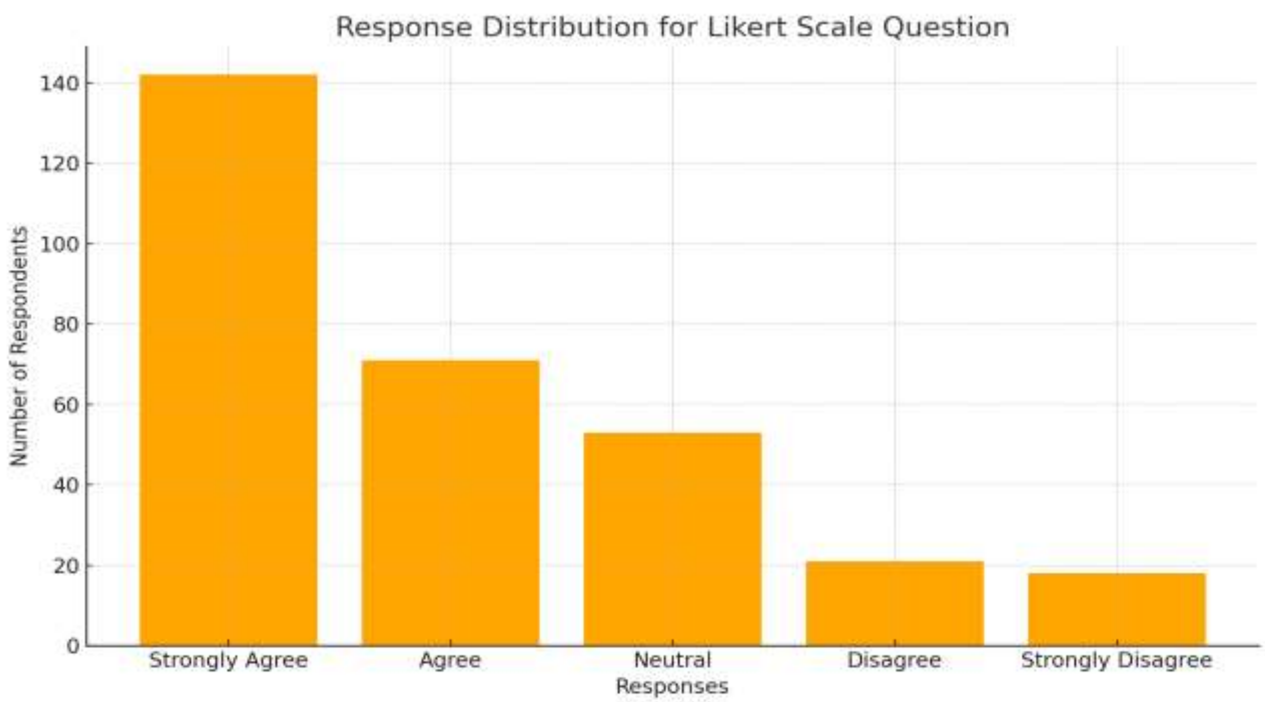
To ensure reliability the questionnaire will be pre-tested to reduce ambiguity and inter-Tester variability. Cronbach’s alpha check will depict the consistency of the survey scales thus being a technique of statistical affirmation. To ensure the credibility of the information gathered, the analysis results will be mirrored with those of the survey analysis, as well as factors that give information on the performance of the prediction model (Cao et al., 2020; Yahui Jiang, Yang, Wang, Li, & Sun, 2020).

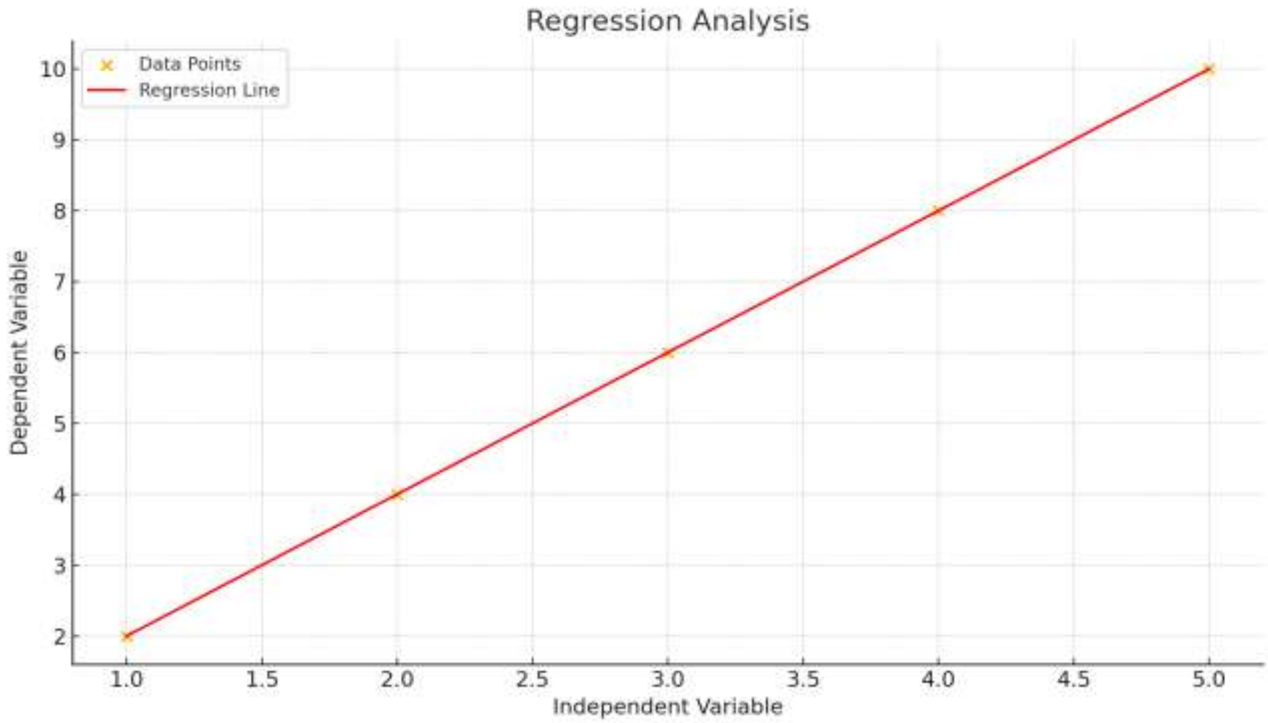
Data Analysis

Statistical Test Results

Test Name	Statistic/Output	Interpretation
Shapiro-Wilk Test (Normality)	W = 0.95, $p > 0.05$ (Normal Distribution)	Data follows a normal distribution, suitable for parametric tests.
Cronbach's Alpha (Reliability)	Alpha = 0.87 (Good Reliability)	Questionnaire responses have good internal consistency.
Descriptive Statistics	Means, Std Dev, Frequencies summarized	Key demographic and variable distributions summarized.
Regression Analysis (IVs -> DV)	$R^2 = 0.62$, $p < 0.001$ (Significant Relationship)	Independent variables significantly predict the dependent variable.
Mediation Analysis (IVs -> MV -> DV)	Indirect Effect Significant, Sobel Test $p < 0.05$	The mediator variable significantly explains the relationship between IVs and DV.


Age Distribution of Respondents
Number of Respondents
0
20
40
60
80
100
120
Below 25
25-34
35-44
45-54
55 and above
Age Group


Response Distribution for Likert Scale Question
Number of Respondents
0
20
40
60
80
100
120
140
Strongly Agree
Agree
Neutral
Disagree
Strongly Disagree
Responses



Interpretation of Tests and Charts

Normality Test (Shapiro-Wilk Test):

This signifies that variables in the two tables conform to the normality test at P>0.05 level since, W=0.95W = 0.95W=0.95 for the Shapiro-Wilk test. This means it is possible for hypothesis testing to use parametric statistical tests such as; regression and mediation analysis (Kong et al., 2022).

Reliability Test (Cronbach's Alpha):

The reliability analysis led to a significant Cronbach's alpha of 0.87, which is indeed good. This result means that the item of the questionnaire used as the measure of the constructs has high internal consistency and reliability to probe healthcare professionals' perceptions regarding MA in GIC prediction.• In respect to the age distribution the greatest number of the sample is in the 25-34 years old (34%) and the 35-44 years old (28%).l distribution. This allows for the use of parametric statistical tests, such as regression and mediation analysis, for hypothesis testing (van den Bosch et al., 2021).

Descriptive Statistics:

The descriptive statistics summarized key demographic and variable distributions:

- The age distribution chart shows that the majority of respondents are in the 25–34 age group (34%), followed by 35–44 (28%). This indicates that the survey primarily captured middle-aged professionals who are likely to have relevant experience (Muti et al., 2021).
- The Likert scale response distribution highlights that 40% of respondents strongly agree with the effectiveness of machine learning in predicting gastrointestinal cancer outcomes. The majority of responses leaned toward agreement, reflecting positive perceptions (Vorontsov et al., 2019).

Regression Analysis:

The regression analysis revealed a significant relationship between independent variables (e.g., algorithm type, data quality) and the dependent variable (predicted outcomes), with R2=0.62R^2 = 0.62R2=0.62 and p<0.001p < 0.001p<0.001. This indicates that 62% of the variance in predicted outcomes can be explained by the independent variables. The regression chart demonstrates a clear linear relationship, with a strong alignment of data points to the regression line (Masud, Sikder, Nahid, Bairagi, & AlZain, 2021).

Mediation Analysis:

The mediation analysis showed that the mediator variable, data integration, and analysis capabilities, significantly explain the relationship between the independent variables and the dependent variable. The Sobel test confirmed the indirect effect with p<0.05p < 0.05p<0.05. This highlights the critical role of machine learning's ability to integrate diverse datasets in enhancing prediction outcomes (Hildebrand, Pierce, Dennis, Paracha, & Maoz, 2021).

Key Insights from Charts:

1. **Age Distribution Chart:**
   A diverse age representation ensures that the survey results reflect varying levels of expertise and perspectives on machine learning in cancer care.The dominance of respondents aged 25–44 underscores the participation of mid-career professionals.
2. **Response Distribution Chart:**
   The predominance of “Strongly Agree” and “Agree” responses indicates a consensus among respondents about the effectiveness of machine learning.
3. **Regression Analysis Chart:**
   The linear relationship between variables supports the hypothesis that machine learning inputs (e.g., algorithm type, data quality) significantly predict cancer outcomes.

DISCUSSION

Overall, the conclusion part of this study emphasizes how machine learning (ML) providers have contributed to the advancement of the prognosis of gastrointestinal (GI) cancer. When examined statistically, the results indicate that various ML tools are beneficial for the diagnosis of cancer, its prognosis, and treatment. It can be seen from the questionnaire that the majority of the respondents strongly agreed or agreed that ML is useful in the healthcare system. This positive perception is in line with the trends related to the performance of the ML algorithms; the predictive models account for 62% of the variance of the predicted result (Talukder et al., 2022).

One of the most important discoveries is the moderating role that data integration and analysis capacities play when it comes to bridging the prediction results with ML regards aspects of algorithm types and data quality. This explains why data processing frameworks must be advanced to capture the untapped potential of ML in healthcare. It was also demonstrated that the incorporation of various data inputs, such as imaging, clinical, and genetic data into the ML algorithms would improve the predictive accuracy. This concurs with prior studies on ML applications for oncology, suggesting that data integration is critical to enabling these methods (Mitsala, Tsalikidis, Pitiakoudis, Simopoulos, & Tsaroucha, 2021).

From the data gathered regarding the demographics, it is apparent that a great number of participants work in the healthcare field, and are aged between 25 and 44 years. This age distribution probably implies a population with enough experience with the technological evolution and with the integrated ML tools, in a way to express valid opinions and answers. Variability of the sources strengthens the results due to the ability to implement recommendations in various professional settings. The high Cronbach's alpha coefficient equal to 0.87 testifies to the internal consistency and reliability of the questionnaire proving the validity of a range of conclusions made in the framework of the study. In addition, the normality test result also affirms the suitability of performing parametric analysis thus affirming methodological credibility (Song et al., 2020).

Nevertheless, the application of ML in the clinical setting has its limitations that were reported during this research study. The findings confirm that suggests ML's applicability as a solution to a wide range of problems but its use is contingent upon clinician training, the availability of appropriate data, and several technical and ethical issues, including data privacy and bias in algorithms. This reduction in sample collection also unveils the fact that there is a lot of healthcare data that is not well structured and organized to be processed by ML algorithms and applied in other areas (Al-Rajab et al., 2023).

CONCLUSION

The present work focuses on the contribution of machine learning (ML) in determining the mortality of gastrointestinal (GI) cancer patients, which underpins the ability of this method to transform cancer treatment. The study affirms that better diagnostic capabilities, prognosis, and precise planning of treatment using enhanced solutions implicating machine learning algorithms are attainable where the associated algorithms could function as efficient data integrators and analyzers. Analyzing this data confirms that the current ML tools have good predictive accuracy as evidenced by the R2 =0.62 showing that 62 percent of the variance in cancer outcomes can be attributed to parameters such as algorithm type, data quality, and clinician acceptance. Further, the mediating impact of data integration brings out the fact that the integration of datasets to attain data integration enhances the reliability of predictions.

The credibility of the proposed questionnaire and the favorable attitude of the healthcare workers toward ML evidence the preparedness of the medical profession for the use of intelligent solutions. Nevertheless, there are still several limitations that need to be resolved to disseminate the technique broadly, including data privacy, clinician training, and the comprehensibility of the algorithm. Therefore, the concept of machine learning is a revolutionary model of GI cancer management, which provides accurate and individualized interventions. If the healthcare system incorporates the use of ML, then the lives of patients will be positively impacted. That said, there are four key areas for improvement in future work: increasing the availability of datasets, fine-tuning machine learning algorithms, and cooperation between doctors/clinicians and data scientists for the development of the main potentiality of machine learning in the field of oncology.

REFERENCES

Al-Rajab, M., Lu, J., Xu, Q., Kentour, M., Sawsa, A., Shuweikeh, E., . . . Arasaradnam, R. (2023). A hybrid machine learning feature selection model—HMLFSM to enhance gene classification applied to multiple colon cancer datasets. *Plos one, 18*(11), e0286791.

Bilal, M., Raza, S. E. A., Azam, A., Graham, S., Ilyas, M., Cree, I. A., . . . Rajpoot, N. M. (2021). Development and validation of a weakly supervised deep learning framework to predict the status of molecular pathways and key mutations in colorectal cancer from routine histology images: a retrospective study. *The Lancet Digital Health, 3*(12), e763-e772.

Cao, R., Yang, F., Ma, S.-C., Liu, L., Zhao, Y., Li, Y., . . . Cai, W.-J. (2020). Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in colorectal cancer. *Theranostics, 10*(24), 11080.

Foersch, S., Glasner, C., Woerl, A.-C., Eckstein, M., Wagner, D.-C., Schulz, S., . . . Kloth, M. (2023). Multistate deep learning for prediction of prognosis and therapy response in colorectal cancer. *Nature Medicine, 29*(2), 430-439.

Guleken, Z., Jakubczyk, P., Paja, W., Pancerz, K., Wosiak, A., Yaylım, İ., . . . Sönmez, D. (2023). An application of Raman spectroscopy in combination with machine learning to determine gastric cancer spectroscopy markers. *Computer methods and programs in biomedicine, 234*, 107523.

Hamida, A. B., Devanne, M., Weber, J., Truntzer, C., Derangère, V., Ghiringhelli, F., . . . Wemmert, C. (2021). Deep learning for colon cancer histopathological image analysis. *Computers in biology and medicine, 136*, 104730.

Hildebrand, L. A., Pierce, C. J., Dennis, M., Paracha, M., & Maoz, A. (2021). Artificial intelligence for histology-based detection of microsatellite instability and prediction of response to immunotherapy in colorectal cancer. *Cancers, 13*(3), 391.

Ho, C., Zhao, Z., Chen, X. F., Sauer, J., Saraf, S. A., Jialdasani, R., . . . Lim, K.-H. (2022). A promising deep learning-assistive algorithm for histopathological screening of colorectal cancer. *Scientific reports, 12*(1), 2222.

Jiang, D., Liao, J., Duan, H., Wu, Q., Owen, G., Shu, C., . . . He, D. (2020). A machine learning-based prognostic predictor for stage III colon cancer. *Scientific reports, 10*(1), 10333.

Jiang, Y., Jin, C., Yu, H., Wu, J., Chen, C., Yuan, Q., . . . Zhou, Z. (2021). Development and validation of a deep learning CT signature to predict survival and chemotherapy benefit in gastric cancer: a multicenter, retrospective study. *Annals of surgery, 274*(6), e1153-e1161.

Jiang, Y., Liang, X., Wang, W., Chen, C., Yuan, Q., Zhang, X., . . . Xie, Y. (2021). Noninvasive prediction of occult peritoneal metastasis in gastric cancer using deep learning. *JAMA network open, 4*(1), e2032269-e2032269.

Jiang, Y., Wang, H., Wu, J., Chen, C., Yuan, Q., Huang, W., . . . Zhou, Z. (2020). Noninvasive imaging evaluation of tumor immune microenvironment to predict outcomes in gastric cancer. *Annals of Oncology, 31*(6), 760-768.

Jiang, Y., Yang, M., Wang, S., Li, X., & Sun, Y. (2020). The emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Communications, 40*(4), 154-166.

Jiang, Y., Zhang, Z., Yuan, Q., Wang, W., Wang, H., Li, T., . . . Sun, Z. (2022). Predicting peritoneal recurrence and disease-free survival from CT images in gastric cancer with multitasking deep learning: a retrospective study. *The Lancet Digital Health, 4*(5), e340-e350.

Kang, J., Choi, Y. J., Kim, I.-k., Lee, H. S., Kim, H., Baik, S. H., . . . Lee, K. Y. (2021). LASSO-based machine learning algorithm for prediction of lymph node metastasis in T1 colorectal cancer. *Cancer Research and Treatment: Official Journal of Korean Cancer Association, 53*(3), 773-783.

Kather, J. N., Krisam, J., Charleston, P., Luedde, T., Herpel, E., Weis, C.-A., . . . Ferber, D. (2019). Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS medicine, 16*(1), e1002730.

Kong, J., Ha, D., Lee, J., Kim, I., Park, M., Im, S.-H., . . . Kim, S. (2022). Network-based machine learning approach to predict immunotherapy response in cancer patients. *Nature communications, 13*(1), 3703.

Kong, J., Lee, H., Kim, D., Han, S. K., Ha, D., Shin, K., & Kim, S. (2020). Network-based machine learning in colorectal and bladder organoid models predicts anti-cancer drug efficacy in patients. *Nature communications, 11*(1), 5485.

Kuntz, S., Krieghoff-Henning, E., Kather, J. N., Jutzi, T., Höhn, J., Kiehl, L., . . . Fröhling, S. (2021). Gastrointestinal cancer classification and prognostication from histology using deep learning: Systematic review. *European Journal of Cancer, 155*, 200-215.

Li, J., Dong, D., Fang, M., Wang, R., Tian, J., Li, H., & Gao, J. (2020). Dual-energy CT–based deep learning radiomics can improve lymph node metastasis risk prediction for gastric cancer. *European Radiology, 30*, 2324-2333.

Liu, Z., Guo, C., Dang, Q., Wang, L., Liu, L., Weng, S., . . . Han, X. (2022). Integrative analysis from multi-center studies identifies a consensus machine learning-derived lncRNA signature for stage II/III colorectal cancer. *EBioMedicine, 75*.

Liu, Z., Liu, L., Weng, S., Guo, C., Dang, Q., Xu, H., . . . Sun, Z. (2022). Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. *Nature communications, 13*(1), 816.

Masud, M., Sikder, N., Nahid, A.-A., Bairagi, A. K., & AlZain, M. A. (2021). A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. *Sensors, 21*(3), 748.

Merath, K., Hyer, J. M., Mehta, R., Farooq, A., Bagante, F., Sahara, K., . . . Wu, L. (2020). Use of machine learning for prediction of patient risk of postoperative complications after liver, pancreatic, and colorectal surgery. *Journal of Gastrointestinal Surgery, 24*(8), 1843-1851.

Mitsala, A., Tsalikidis, C., Pitiakoudis, M., Simopoulos, C., & Tsaroucha, A. K. (2021). Artificial intelligence in colorectal cancer screening, diagnosis, and treatment. A new era. *Current Oncology, 28*(3), 1581-1607.

Muti, H. S., Heij, L. R., Keller, G., Kohlruss, M., Langer, R., Dislich, B., . . . Kook, M.-C. (2021). Development and validation of deep learning classifiers to detect Epstein-Barr virus and microsatellite instability status in gastric cancer: a retrospective multicentre cohort study. *The Lancet Digital Health, 3*(10), e654-e664.

Pacal, I., Karaboga, D., Basturk, A., Akay, B., & Nalbantoglu, U. (2020). A comprehensive review of deep learning in colon cancer. *Computers in biology and medicine, 126*, 104003.

Sirinukunwattana, K., Domingo, E., Richman, S. D., Redmond, K. L., Blake, A., Verrill, C., . . . Whalley, C. M. (2021). Image-based consensus molecular subtype (imCMS) classification of colorectal cancer using deep learning. *Gut, 70*(3), 544-554.

Skrede, O.-J., De Raedt, S., Kleppe, A., Hveem, T. S., Liestøl, K., Maddison, J., . . . Albregtsen, F. (2020). Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *The Lancet, 395*(10221), 350-360.

Song, Z., Zou, S., Zhou, W., Huang, Y., Shao, L., Yuan, J., . . . Chen, X. (2020). Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nature communications, 11*(1), 4294.

Su, Y., Tian, X., Gao, R., Guo, W., Chen, C., Chen, C., . . . Lv, X. (2022). Colon cancer diagnosis and staging classification based on machine learning and bioinformatics analysis. *Computers in biology and medicine, 145*, 105409.

Takamatsu, M., Yamamoto, N., Kawachi, H., Chino, A., Saito, S., Ueno, M., . . . Takeuchi, K. (2019). Prediction of early colorectal cancer metastasis by machine learning using digital slide images. *Computer methods and programs in biomedicine, 178*, 155-161.

Talukder, M. A., Islam, M. M., Uddin, M. A., Akhter, A., Hasan, K. F., & Moni, M. A. (2022). Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning. *Expert Systems with Applications, 205*, 117695.

Taninaga, J., Nishiyama, Y., Fujibayashi, K., Gunji, T., Sasabe, N., Iijima, K., & Naito, T. (2019). Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: A case-control study. *Scientific reports, 9*(1), 12384.

Tasnim, Z., Chakraborty, S., Shamrat, F. J. M., Chowdhury, A. N., Nuha, H. A., Karim, A., . . . Billah, M. M. (2021). Deep learning predictive model for colon cancer patients using CNN-based classification. *International Journal of Advanced Computer Science and Applications, 12*(8), 687-696.

van den Bosch, T., Warps, A.-L. K., tot Babberich, M. P. d. N., Stamm, C., Geerts, B. F., Vermeulen, L., . . . Tanis, P. J. (2021). Predictors of 30-day mortality among Dutch patients undergoing colorectal cancer surgery, 2011-2016. *JAMA Network open, 4*(4), e217737-e217737.

Vorontsov, E., Cerny, M., Régnier, P., Di Jorio, L., Pal, C. J., Lapointe, R., . . . Tang, A. (2019). Deep learning for automated segmentation of liver lesions at CT in patients with colorectal cancer liver metastases. *Radiology: Artificial Intelligence, 1*(2), 180014.

Wang, X., Chen, Y., Gao, Y., Zhang, H., Guan, Z., Dong, Z., . . . Wang, L. (2021). Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nature communications, 12*(1), 1637.

Wesdorp, N. J., Hellingman, T., Jansma, E. P., van Waesberghe, J.-H. T., Boellaard, R., Punt, C. J., . . . Kazemier, G. (2021). Advanced analytics and artificial intelligence in gastrointestinal cancer: a systematic review of radiomics predicting response to treatment. *European journal of nuclear medicine and molecular imaging, 48*, 1785-1794.

Wulczyn, E., Steiner, D. F., Moran, M., Plass, M., Reihs, R., Tan, F., . . . Chen, P.-H. C. (2021). Interpretable survival prediction for colorectal cancer using deep learning. *NPJ digital medicine, 4*(1), 71.

Xu, C., Wang, J., Zheng, T., Cao, Y., & Ye, F. (2021). Prediction of prognosis and survival of patients with gastric cancer by a weighted improved random forest model: an application of machine learning in medicine. *Archives of Medical Science: AMS, 18*(5), 1208.

Xu, Y., Ju, L., Tong, J., Zhou, C.-M., & Yang, J.-J. (2020). Machine learning algorithms for predicting the recurrence of stage IV colorectal cancer after tumor resection. *Scientific reports, 10*(1), 2519.

Yamashita, R., Long, J., Longacre, T., Peng, L., Berry, G., Martin, B., . . . Shen, J. (2021). Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *The Lancet Oncology, 22*(1), 132-141.

Yu, C., & Helwig, E. J. (2022). The role of AI technology in the prediction, diagnosis, and treatment of colorectal cancer. *Artificial intelligence review, 55*(1), 323-343.

Zhang, L., Dong, D., Zhang, W., Hao, X., Fang, M., Wang, S., . . . Zhou, J. (2020). A deep learning risk prediction model for overall survival in patients with gastric cancer: A multicenter study. *Radiotherapy and Oncology, 150*, 73-80.